Reviewer's report

Title: Perceived neighborhood problems: multilevel analysis with ecometric properties and socioeconomic associated factors in a Southern adult Brazilian population

Version: 1 Date: 23 April 2013

Reviewer: Daniel J J Corsi

Reviewer's report:

This study examined the ecometric and psychometric properties of adults residents’ perceptions of neighborhood problems AND the association of neighborhood problems with individual and census tract characteristics. I have several suggestions that may improve the quality of this manuscript. These are outlined below.

Major Revisions:

1. Scope of the paper: This appears to be a methods paper. The authors have described an objective to examine the ecometric and psychometric properties of the instrument. However they have also assessed the relationship between individual and neighborhood level characteristics and the neighborhood problem scales described. I would be inclined to re-frame the second objective as an extension of the first objective where you are assessing the predictive validity of the instrument. Also this needs a theoretical discussion: for example why would you expect neighborhood problems to vary by the income level or other characteristics of neighborhoods?

1. I would like to see a conceptual framework in the introduction. What are the pathways where by the perceived neighborhood problems are likely to influence health. What advantages do the perception-based measures have over objective-based ones in the area of neighborhood problems? And how do your proposed scales capture these constructs?

The differences in the objective and perception based methodologies you have described are well known and there are many studies which are using similar approaches to what you have described some of which are ongoing in low and middle income countries including Brazil. See for example: Corsi DJ et al. Environmental Profile of a Community's Health (EPOCH): an ecometric assessment of measures of the community environment based on individual perception. PLoS One. 2012;7(9):e44410. doi: 10.1371/journal.pone.0044410 AND

Finally, what particular health outcomes do you anticipate will be associated with such perceptions?

Methods:

2. Give further details about the EpiFloripa study. What are its objectives and measures? How long will the study continue for? Why were only adults 20-59 included?

3. It is not clear that you have used the principal factor analysis in order to group the 16 items into 2 scales (this comes out in the results). Again you need to tie this back to the conceptual framework. Did the items group onto the scales in the expected manner? Was the set of questions (adapted from a UK study) appropriate for the Brazilian context?

4. Ecometric properties: The reliability is a function of the item inconsistency, the between-individual differentials in agreement, and the between-neighborhood variation. Therefore this measure will also increase as the number of individuals within neighborhoods increase and the number of scale items increases.

5. The authors should provide more details on how they estimated the multilevel models (ML, RML, Bayesian ...?). For example, it looks like you used Bayesian however the reference (page 9) to “Bayes estimates” is rather cryptic. Please also describe the use of weights in the ML models. Why was this necessary? Did the results change versus the non weighted estimates?

6. There needs to be further description of the second objective which is currently missing from the methods section. As noted previously, it would be interesting to frame this objective in the sense of ‘predictive’ validity of your measures. Therefore you should state your then analytical approach to answer your hypothesis here (e.g. that the low income neighborhoods would have more problems). Also how did you decide which neighborhood-level characteristic to use? Where other measures available, for example % in poverty, % of different racial/ethnic groups?

8. Explain how you developed the single composite neighborhood level scores to be used in these models. Did this come from the multilevel analysis or some other aggregation procedure?

7. Results: I would re-organize into the reliability and instrument properties component and then the predictive validity.

8. The scales looked somewhat correlated. Did you fit any models including both scales on the right hand side? How did this change the results?

9. The reliability and ICC statistics in this study seem very robust. This leads me to believe that you have more than adequate number of respondents and
questionnaire items. Have you tried to examine the sensitivity of your findings by eliminating certain questionnaire items - for example those with low factor loadings? If you could demonstrate similarly high reliability with fewer items this would be important in future studies to reduce respondent burden.

10. Discussion: I would like to see some discussion of how these findings will be used in the Epifloripa study. Do you plan to use these scales as explanatory variables in regression models to examine health outcomes?

11. Table 5 is not clear. Did you fit 3 separate models for each of the scales? If so then you have not displayed the coefficients for all models. For example how did the coefficients for sex, age, length of residence time change when adding socioeconomic characteristics? and how did these change again when adding the neighborhood characteristic?

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests